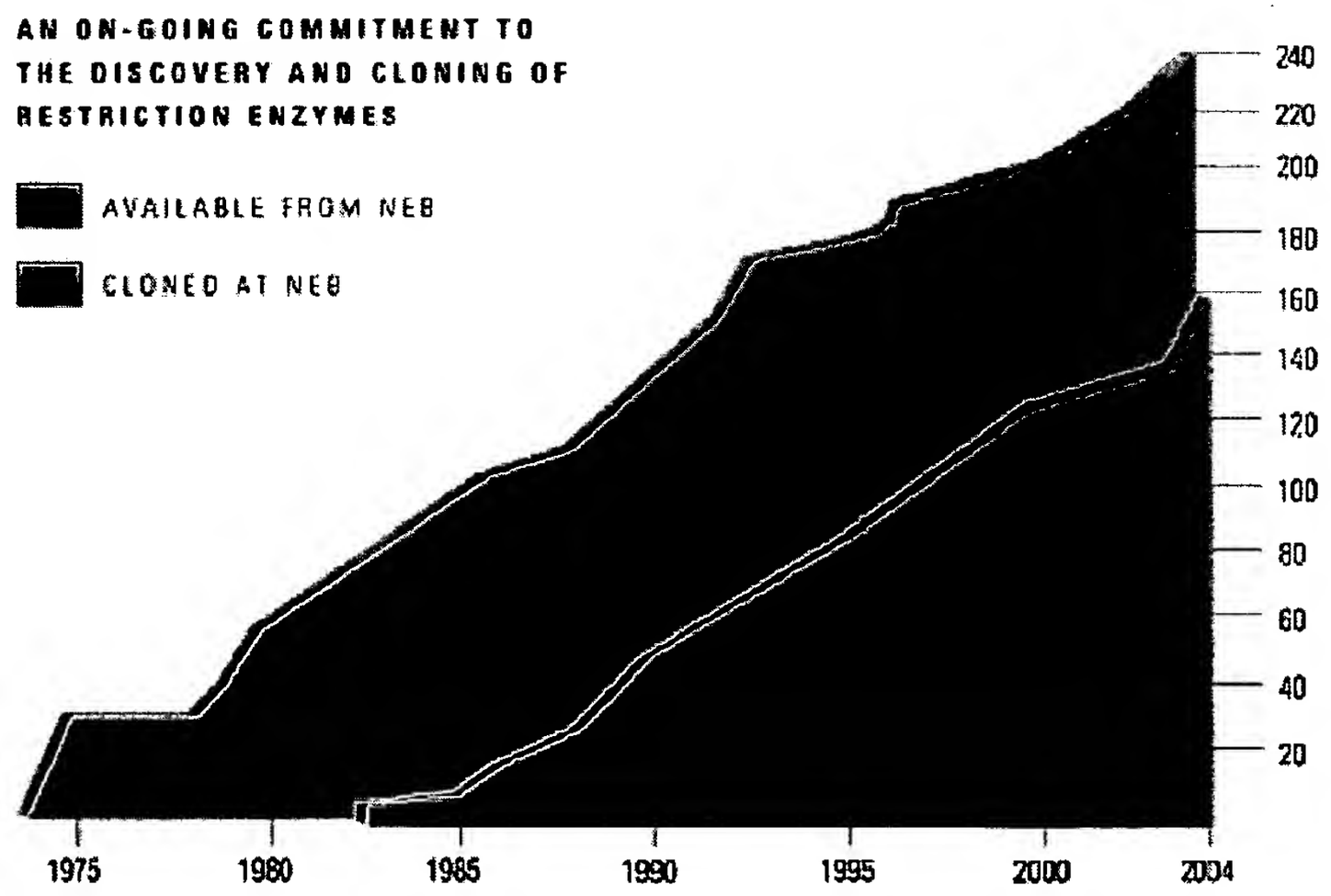
Restriction Endonucleases Overview

Restriction enzymes were discovered about 30 years ago during investigations into the phenomenon of host-specific restriction and modification of bacterial viruses. Bacteria initially resist infections by new viruses, and this "restriction" of viral growth stemmed from endonucleases within the cells that destroy foreign DNA molecules. Among the first of these "restriction enzymes" to be purified were EcoRI and EcoRII from *E.coli*, and HindII and HindIII from Haemophilus influenzae. These enzymes were found to cleave DNA at specific sites, generating discrete, gene-size fragments that could be re-joined in the laboratory. Researchers were quick to recognize that restriction enzymes provided them with a remarkable new tool for investigating gene organization, function and expression.

As the use of restriction enzymes spread among molecular biologists in the late 1970's, companies such as New England Biolabs began to search for more. Except for certain viruses, restriction enzymes were found only within prokaryotes. Many thousands of bacteria and archae have now been screened for their presence. Analysis of sequenced prokaryotic genomes indicates that they are common--all free-living bacteria and archaea appear to code for them.

Restriction enzymes are exceedingly varied; they range in size from the diminutive PvuII (157 amino acids) to the giant CjeI (1250 amino acids) and beyond. Among over 3,000 activities that have been purified and characterized, more than 250 different sequence-specificities have been discovered. Of these, over 30% were discovered and characterized at New England Biolabs. The search for new specificities continues, both biochemically, by the analysis of cell-extracts, and computationally, by the analysis of sequenced genomes. Although most activities encountered today turn out to be duplicates--isoschizomers--of existing specificities, restriction enzymes with new specificities are found with regularity.

Exhibit A



Beginning in the early 1980's, New England Biolabs embarked on a program to clone and overexpress the genes for restriction enzymes. Cloning improves enzyme purity by separating enzymes from contaminating activities present in the same cells. It also improves enzyme yields and greatly simplifies purification, and it provides the genes for sequencing and analysis, and the proteins for x-ray crystallography.

Restriction enzymes protect bacteria from infections by viruses, and it is generally accepted that this is their role in nature. They function as microbial immune systems. When a strain of E.coli lacking a restriction enzyme is infected with a virus, most virus particles can initiate a successful infection. When the same strain contains a restriction enzyme, however, the probability of successful infection plummets. The presence of additional enzymes has a multiplicative effect; a cell with four or five independent restriction enzymes could be virtually impregnable.

Restriction enzymes usually occur in combination with one or two modification enzymes (DNA-methyltransferases) that protect the cell's own DNA from cleavage by the restriction enzyme. Modification enzymes recognize the same DNA sequence as the restriction enzyme that they accompany, but instead of cleaving the sequence, they methylate one of the bases in each of the DNA strands. The methyl groups protrude into the major groove of DNA at the binding site and prevent the restriction enzyme from acting upon it. Together, a restriction enzyme and its "cognate" modification enzyme(s) form a restriction-modification (R-M) system. In some R-M systems the restriction enzyme and the modification enzyme(s) are separate proteins that act independently of each other. In other systems, the two activities occur as separate subunits, or as separate domains, of a larger, combined, restriction-and-modification enzyme.

Restriction enzymes are traditionally classified into three types on the basis of subunit composition, cleavage position, sequence-specificity and cofactor-requirements. However, amino acid sequencing has uncovered extraordinary variety among restriction enzymes and revealed that at the molecular level there are many more than three different kinds.

Type I enzymes are complex, multisubunit, combination restriction-and-modification enzymes that cut DNA at random far from their recognition sequences. Originally thought to be rare, we now know from the analysis of sequenced genomes that they are common. Type I enzymes are of considerable biochemical interest but they have little practical value since they do not produce discrete restriction fragments or distinct gel-banding patterns.

Type II enzymes cut DNA at defined positions close to or within their recognition sequences. They produce discrete restriction fragments and distinct gel banding patterns, and they are the only class used in the laboratory for DNA analysis and gene cloning. Rather then forming a single family of related proteins, type II enzymes are a collection of unrelated proteins of many different sorts. Type II enzymes frequently differ so utterly in amino acid sequence from one another, and indeed from every other known protein, that they likely arose independently in the course of evolution rather than diverging from common ancestors.

The most common type II enzymes are those like HhaI, HindIII and NotI that cleave DNA within their recognition sequences. Enzymes of this kind are the principle ones available commercially. Most recognize DNA sequences that are symmetric because they bind to DNA as homodimers, but a few, (e.g., BbvCI: CCTCAGC) recognize asymmetric DNA sequences because they bind as heterodimers. Some enzymes recognize continuous sequences (e.g., EcoRI: GAATTC) in which the two half-sites of the recognition sequence are adjacent, while others recognize discontinuous sequences (e.g., BglI: GCCNNNNNGGC) in which the half-sites are separated. Cleavage leaves a 3′-hydroxyl on one side of each cut and a 5′-phosphate on the other. They require only magnesium for activity and the corresponding modification enzymes require only S-adenosylmethionine. They tend to be small, with subunits in the 200–350 amino acid range.

The next most common type II enzymes, usually referred to as 'type IIs" are those like FokI and AlwI that cleave outside of their recognition sequence to one side. These enzymes are intermediate in size, 400–650 amino acids in length, and they recognize sequences that are continuous and asymmetric. They comprise two distinct domains, one for DNA binding, the other for DNA cleavage. They are thought to bind to DNA as monomers for the most part, but to cleave DNA cooperatively, through dimerization of the cleavage domains of adjacent enzyme molecules. For this reason, some type IIs enzymes are much more active on DNA molecules that contain multiple recognition sites.

The third major kind of type II enzyme, more properly referred to as "type IV" are large, combination restriction-and-modification enzymes, 850–1250 amino acids in length, in which the two enzymatic activities reside in the same protein chain. These enzymes cleave outside of their recognition sequences; those that recognize continuous sequences (e.g., Eco57I: CTGAAG) cleave on just one side; those that recognize discontinuous sequences (e.g., BcgI: CGANNNNNNTGC) cleave on both sides releasing a small fragment containing the recognition sequence. The amino acid sequences of these enzymes are varied but their organization are consistent. They comprise an N-terminal DNA-cleavage domain joined to a DNA-modification domain and one or two DNA sequence-specificity domains forming the C-terminus, or present as a separate subunit. When these enzymes bind to their substrates, they switch into either restriction mode to cleave the DNA, or modification mode to methylate it.

Type III enzymes are also large combination restriction-and-modification enzymes. They cleave outside of their recognition sequences and require two such sequences in opposite orientations within the same DNA molecule to accomplish cleavage; they rarely give complete digests. No laboratory uses have been devised for them, and none are available commercially.